Defense Threat Reduction Agency
8725 John J. Kingman Road, MS
6201 Fort Belvoir, VA 22060-6201

DTRA-TR-11-20

TECHNICAL REPORT

A Small-World Network Model of Disease Transmission

Approved for public release; distribution is unlimited

December 2011

HDTRA1-07-C-0066

Karen E. Cheng, et al.

Prepared by:
Applied Research Associates, Inc.
Health Effects and Medical Response Group
801 N. Quincy St.
Arlington, VA 22203

DESTRUCTION NOTICE:

Destroy this report when it is no longer needed.
Do not return to sender.

PLEASE NOTIFY THE DEFENSE THREAT REDUCTION AGENCY, ATTN: DTRIAC OP-ONIU, 8725 JOHN J. KINGMAN ROAD, MS-6201, FT BELVOIR, VA 22060-6201, IF YOUR ADDRESS IS INCORRECT, IF YOU WISH THAT IT BE DELETED FROM THE DISTRIBUTION LIST, OR IF THE ADDRESSEE IS NO LONGER EMPLOYED BY YOUR ORGANIZATION.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
00-12-2011	Technical Report	May 2009 - May 2011

4. TITLE AND SUBTITLE
A Small-World Network Model of Disease Transmission

5a. CONTRACT NUMBER
HDTRA1-07-C-0066

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Cheng, K.
Crary, D.
Rodriguez, J.

5d. PROJECT NUMBER
JSTO Project CB07MSB100

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Applied Research Associates, Inc.
801 N. Quincy St. Ste 600
Arlington, VA 22203

8. PERFORMING ORGANIZATION REPORT NUMBER
ARA-TR-10-SEASSP-17921-003

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Threat Reduction Agency
8725 John J. Kingman Road, MSC
Fort Belvoir, VA 22060

10. SPONSOR/MONITOR'S ACRONYM(S)
DTRA/RD-CBI

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DTRA-TR-11-20

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The most well-known mathematical models for the investigation of the spread of transmissible infections are compartmental models. These compartments categorize the population according to disease status, susceptibility to infection, etc. and by using coupled differential equations the models can often be solved analytically. But these models assume uniform mixing, i.e. any individual in the population has the same probability of contacting any other individual. These models require very detailed contact networks, but the large data set necessary to run the model is usually incomplete. Recently two advances have been made: one incorporates realistic assumptions about transportation and demographics as they affect person-to-person contact. The second has provided tools for describing complex networks and understanding their dynamics. Small-world networks show a small average distance between nodes (measured as the least number of connections) compared with the size of the graph. These nodes can represent either persons or locations. In this paper complicated and unrealistic modeling based on uniform mixing is replaced by a simpler, less computationally extensive, and more realistic small world network model. With this model outbreaks in various geographical locations can be rapidly characterized, analyzed, and preventive measures for outbreak control recognized and recommended.

15. SUBJECT TERMS
Complex networks; disease transmission; secondary infection; epidemic modeling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UNCLASSIFIED					Dr. Christopher Kiley
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	33	19b. TELEPHONE NUMBER *(include area code)*
U	U	U			703-767-3460

Preface

The model of disease transmission in a large population described in this technical report was developed for Project CB07MSB100 of the Joint Science and Technology Office (JSTO) of the Department of Defense (DoD) Chemical and Biological Defense (CBD) Program. JSTO is also the Chemical/Biological Technologies (CB) Directorate in the Research and Development (RD) Enterprise of the Defense Threat Reduction Agency (DTRA). Project CB07MSB100 is titled *Predicting Effects Due to Infectious/Contagious Diseases for JEM.*

This project was initiated by Mr. Charles Fromer and Mr. Richard (Rick) Fry of the Information Systems Capability Development Division (RD-CBI). It was funded under DTRA Contract Number HDTRA1-07-C-0066 to Applied Research Associates, Inc. (ARA), with subcontractor Sandia National Laboratories (Sandia). Mr. Fry was the first Contractor Officer's Representative (COR) for this contract. On 28 October 2008, Mrs. Stephanie Hamilton of RD-CBI became the COR and on 10 July 2009, Dr. Christopher Kiley, also of RD-CBI, was named COR for the remainder of the contract. The target application for the product of this contract is the Joint Effects Model (JEM) under the auspices of the Joint Project Manager of Information Systems (JPM IS) of the Joint Program Executive Office for Chemical and Biological Defense (JPEO-CBD).

The integration of the model described in this technical report into a Contagion Model for JEM is described in the final report for this contract.

Table of Contents

Table of Figures

1. Component Overview

The secondary infection small-world network model is written using object-oriented principles. Modelers who add new disease models into the secondary infection model collection should "extend" the Epidemic class and implement the required interfaces.

Programmers who use the Secondary Infection small world network models in their programs will interface with the models via the Secondary Infection SWN.DLL.

The current secondary infection model collection includes three disease models: influenza, plague, and smallpox.

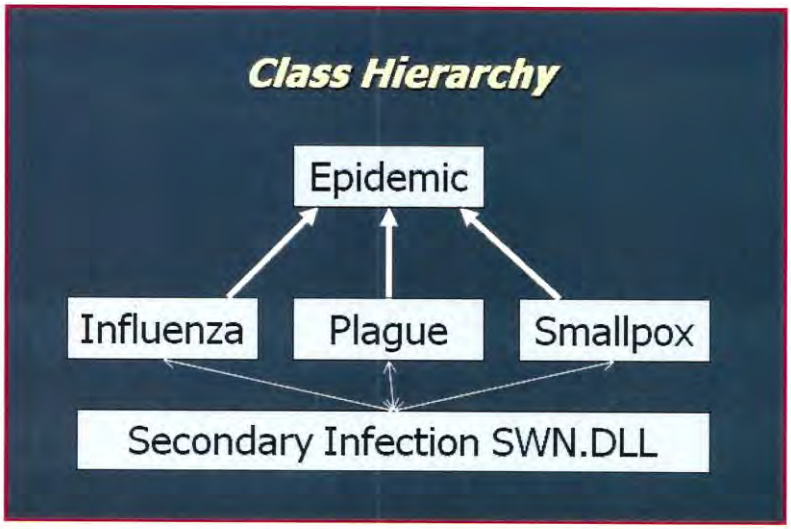


Figure 1-1. Class hierarchy for the secondary infection models

- Authors: Ms. Karen Cheng, Dr. David Crary, Dr. Darren R. Oldson
- Contact: Ms. Karen Cheng, kcheng@ara.com, 703-816-8886 x 138
- Company: Applied Research Associates, Inc.
- Creation Date: February 28, 2010

Given basic properties known about the disease, the population (the number of susceptible individuals, and the number initially infected), the Secondary Infection SWN.dll estimates the number of people daily who are in the following compartments:

- Susceptible [S]
- Exposed [E]
- Infectious [I]
- Removed [R]

It additionally has routines to return the following:

- Maximum day of the outbreak
- The duration of the outbreak
- The number of total infections throughout the outbreak

1.1. Introduction

Mathematical models for the investigation of the spread of infectious disease have a long history [1]. The most well known of these models are the so-called compartmental models, which are identified by the acronyms that indicate the compartments included in the model, such as SIR (susceptible-infectious-recovered), SIS (susceptible-infectious-susceptible), SEIR (susceptible-exposed-infectious-recovered) and variants of these. These models can be described by coupled differential equations, with the number of equations equal to the number of unique compartments in the model, which can be solved analytically in certain cases [2,3], and also have been well studied using stochastic simulation. These models have traditionally been solved under idealized hypotheses. The most important of these hypotheses, for this study, is the assumption of uniform mixing; that is, any individual in the population under study have the same probability of contacting any other individual.

More recently, two related advances have been made which have impacted the field of epidemic modeling. The first of these is the development by Eubank *et al.* [4] of large scale simulations that use realistic estimates of population mobility, based on transportation, census and land use data. From these data, a realistic, dynamic contact graph is created that represents the person-to-person contact network. This contact network is constrained by realistic assumptions about the transportation infrastructure, as well as demographics consistent with distributions in census data. The population mixing implied by this contact graph replaces the uniform mixing assumption of the traditional approach.

Secondly, advances in the experimental and theoretical understanding of network structures that describe a wide range of systems in nature (such as social networks, disease transmissions, cellular chemistry networks, internet connections, world-wide-web structure, phone call networks) has provided new understanding of the classification of complex networks, understanding of their dynamics, and new tools for network diagnostics. In particular, advances in computing speed and computerized data acquisition have provided researchers with large datasets with which to investigate the properties of complex, real-world networks.

As a result of this theoretical research, it has become evident that the vast majority of network structures encountered in real data are one of two types, scale-free and small-world. Referring to the components of a graph as "nodes" and "connections" (or "edges"), scale-free networks show a power-law distribution of connections per node. Small-world networks, on the other hand, show an anomalously small average distance between nodes (measured as the least number of connections required to traverse from one node to another) compared with the size of the graph.

The realistic networks developed by Eubank *et al.* show both scale-free and small world properties. The fundamental structure used by Eubank *et al.* in this investigation is a bipartite graph (a graph with two types of nodes); in this case the nodes are people and places.

This bipartite graph can be projected onto two different contact graphs, which are easier to interpret in terms of the relationships between persons or places than the original bipartite graph. One of these contact graphs is person-centric, where nodes represent persons and edges represent contacts (in

space and time) between those persons. The second of these graphs is location-centric, where nodes represent locations, and edges represent a person traveling between those two locations during the course of the simulation. It is remarkable that in the simulation of Eubank *et al.* the person-centric and location-centric graphs show different well-defined structures, with the person-centric graphs having small-world network structure, and the location graph well represented by a scale-free network.

It is the observation that a realistic contact network of disease spread shows small-world structure, that provides the motivation for this study. In particular, it is natural to ask if it is possible to replace the complicated, computationally expensive simulation of Eubank *et al.* with a simpler model, based on a small world network, which reproduces the essential features of the full simulation without the computational expense, cost and complexity of deriving the network for each population studied? Moreover, is it possible to derive equations that allow the computation of the simpler model's parameters, which would provide an inexpensive and fast way to model outbreaks in various geographical locations, for which no detailed model exists? This capability would solve one of the main problems encountered in using a detailed epidemic models based on realistic contact networks, namely, that model can only be used to describe an outbreak where detailed data on the contact network exist.

This paper describes our methodology for our approach, and our results show that the answers to these questions appear to be "yes". The features of this approach are described in the sections below.

The Secondary Infection software is structured as a software library SWN.dll that provides the following outputs:

- Estimates the number of people in the following compartments on a daily basis:
  - Susceptible [S]
  - Exposed [E]
  - Infectious [I]
  - Removed [R]

The library also gives the following high level summaries of an outbreak:

- Maximum day of the outbreak
- The duration of the outbreak
- The number of total infections throughout the outbreak

2. Definitions, Acronyms, and Abbreviations

- ***Susceptible.*** Susceptible to infection by the particular disease in the study.
- ***Exposed.*** Exposed to the particular disease in the study; infected but not yet infectious.
- ***Infectious.*** Infected by the particular disease in the study and able to infect others.
- ***Removed.*** Either recovered or died from the disease in the study; not able to be infected or to infect others.

3. Engineering Methodology

3.1. Background on Small World Networks

As discussed above, in an abstract sense, networks (such as the social networks) can be represented as a graph of "nodes" and "connections" (or "edges"). For social networks of interest here, the nodes represent people. The connections between adjacent nodes (nodes separated by one edge) represent contacts between people.

The small world network models two types of contacts, 'short range', such as the contact between members of a family, or neighbors, and long range contacts such as occur in random encounters between people. More specifically, the 'small world' property refers to a network with a small number of shortcuts introduced in an otherwise regular, local underlying network structure. This considerably reduces the average 'distance' between any two nodes in a network, where here distance refers to the shortest path between two nodes, measured by the number of connections traversed. In a relationship network of people, for example, small world networks capture the well-known phenomena of strangers being linked by a web of mutual contact. Most people are only acquainted with a small number of other people, but most strangers can be linked though a small number of mutual acquaintances. This phenomenon has become well known recently through the widespread use of social networking sites, such as Facebook and LinkedIn, but the idea dates back to a 1929 hypothesis by the Hungarian author Frigyes Karinthy.

We originally implemented an SIR (Susceptible, Infectious, Removed) model on a network similar to the work done by Saramaki and Kaski [5] (disease models are discussed in more detail in Section 3.2). In this case, the underlying network was a ring, as shown in Figure 3-1. We refer to the number of local network connections in the underlying network by the letter Z, in this case Z=2. A long range connection is shown by a line traversing the center region of the ring. Later we extended this model to use square local network with Z=4, as shown in Figure 3-2. This value for Z is plausible, since the average family size in the U.S. is 3.14 (U.S. Census Bureau: factfinder.census.gov). The extended model also includes an Exposed [E] category.

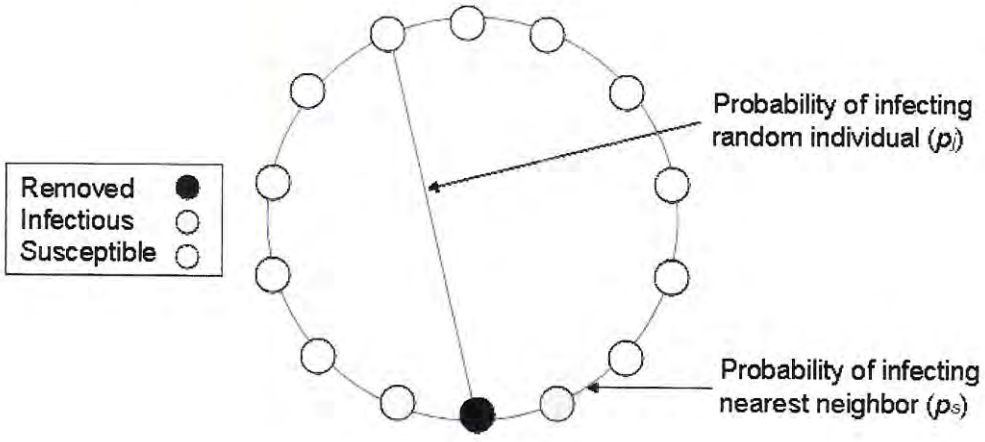


Figure 3-1. SIR Small World Network with Z = 2

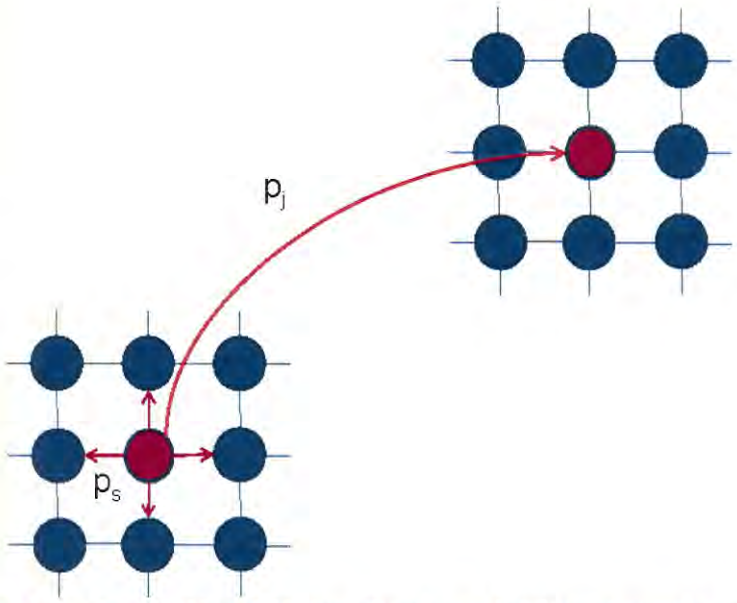


Figure 3-2. Small World Network where lattice parameter Z = 4

Saramaki and Kaski derived a set of coupled differential equations and solved analytically under certain assumptions to give a solution to the number of persons in each of the S, E, I and R compartments as a function of time. We have implemented the same model as a stochastic simulation. The benefits of using this type of simulation are:

- We do not need to make simplifying assumptions of the probabilities p_s and N (the susceptible population) which were required for the analytic solution of Saramaki and Kaski.
- We can easily implement countermeasures in our simulations to study their effects.

A major drawback of large-scale simulations is that they require a long time to run. In our case, however, because the model is simple, it does not require long run times. There is a stochastic nature to the runs but our studies have found that there is not significant variance between the runs (e.g., a small number of runs produce a good representative average).

3.2. Small-World Network Methodology.

As discussed above, the epidemic model used in this study is an SEIR model. In an SEIR each individual in the population at risk is in one of four "compartments" during the progression of the disease. These compartments are specified by the letters SEIR, where S is susceptible (that is, not exposed to the disease), E is exposed but not infectious, I is infectious, and R is removed (either died or recovered).

In SEIR models, an individual, once exposed, moves between these compartments according to certain probability distributions. In fact, the specification of the contact between people producing new exposures, plus the specification of the probability with which an individual moves between compartment after exposure, provides a full specification of the model. In our case, the contact between people is modeled with a regular network where the lattice parameter, Z = 4 (see Figure 3-3). The way individuals move between compartments is specified by an algorithm that involves two parameters for disease transmission, p_s and p_j, where p_s = probability of infecting a nearest neighbor and p_j = probability of infecting a random individual.

The algorithm is described as follows: at every time step of duration Δt, every infectious individual in the network:

1. Infects its nearest neighbors, if susceptible, with probability p_s per neighbor.
    a. When infected, the individual enters the Exposed category and remains there on for a period of time (specified by the *mean and standard deviation of length of time exposed)*. Once this time period is up, the individual enters the Infectious category.
2. With probability p_j, tries to infect one randomly chosen individual, succeeding if the individual is susceptible.
    a. When infected, the randomly chosen individual enters the Exposed category and remains there on for a period of time (specified by the *mean and standard deviation of length of time exposed)*. Once this time period is up, the individual enters the Infectious category.
3. With probability p_r recovers and can no longer be infected or infect others.
    a. The individual enters the Recovered category.

3.3. Advantages of approach

Traditional models of epidemics rely on deriving transmission rates from historical data. There are various limitations to this approach. Embedded in this transmission rates are the original conditions (population size, initial number of infected) and conditions that occurred as the epidemic progressed (containment measures, prophylaxis) and it is unclear whether these historical transmission rates are valid in other contexts. The small world network model allows us to take advantage of the models obtained from realistic complex simulations such as EpiSims [4]. The problem with complex simulations is typically they are time-consuming to run, and obtaining the data for other cities is costly.

A large advantage of the small world network is that it naturally scales the results to different initial conditions such as different number of initially infected, or different population size.

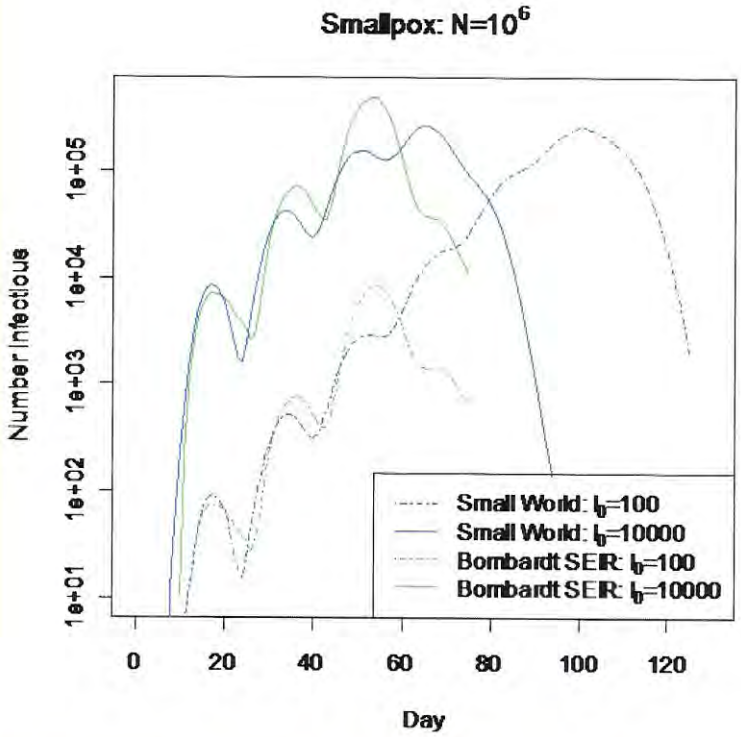


Figure 3-3. The small world network scales the epidemic duration realistically relative for different values of I_o.

This scaling is shown in Figure 3-3 above. The small-world network model is able to scale the duration of an outbreak for different values of the number of initial infected, I_o. A higher I_o value should result in an earlier epidemic. The Bombardt SEIR model, currently used in AMedP-8, cannot scale the duration of the outbreak because the formulation is based on transmission functions derived from the original outbreak (the outbreak may only increase marginally in duration for an increased I_o due to the consequence that there are more people coming out of the *R* compartment). The heights in these plots should not be compared because the small world network plots were fit to EpiSims which have no outbreak mitigation efforts, whereas the Bombardt model is fit to a historical outbreak where there were likely to be outbreak containment efforts, such as quarantine.

By studying how the small world network disease model changes with population density, and thereby creating a formula for how the parameters p_s and p_j scale with population density, we are theoretically able to apply the model created by EpiSims [4] for Portland to other cities. The assumption underlying this capability is that disease behavior itself does not change from city to city, it is the population characteristics that do. This assumption (that disease models remain consistent) is also used in the majority of disease modeling tools, such as DoD's Hazard Prediction and Assessment Capability (HPAC).

Because our model is a simulation, it is easy for us to model tradeoffs of preventive measures.

3.4. Model methodology

Our model depends on eight parameters. They are:

Population specific: N, I_0

N= *total number in population*

I_0= *number initially infected*

Disease specific: μ_{ei}, σ_{ei}, μ_{ir}, σ_{ir}

μ_{ei}, σ_{ei} = *Mean and standard deviation of length of time exposed*

μ_{ir}, σ_{ir} = *Mean and standard deviation of length of time infectious*

Network specific: p_s, p_j

p_s= *Probability per day of spread to nearest neighbor*

p_j= *Probability per day of 'jump' to non-neighbor*

The time development of an epidemic *for a specific disease* is known once I_0, p_s, and p_j are known.

Therefore there are two modes to our model, a data fitting mode, and a disease forecasting mode.

The data fitting model, as described in 3.5 below, is used by disease model developers, whereas the forecasting mode is used by programs such as the Joint Effects Model (JEM) to forecast the number of people who will become ill from a particular disease.

3.5. Selection of disease parameters

We had Sandia National Laboratories perform EpiSims [4] runs of influenza, plague, and anthrax for eight different locations in Portland Oregon, for two different numbers of initially infected, 100 and 1000. Fortunately, Portland has many interesting geographical features such as varied population densities and a major river that runs through the city. These eight points were chosen above and below the river, and in different local population densities. Figure 3-4 shows the points chosen for our study.

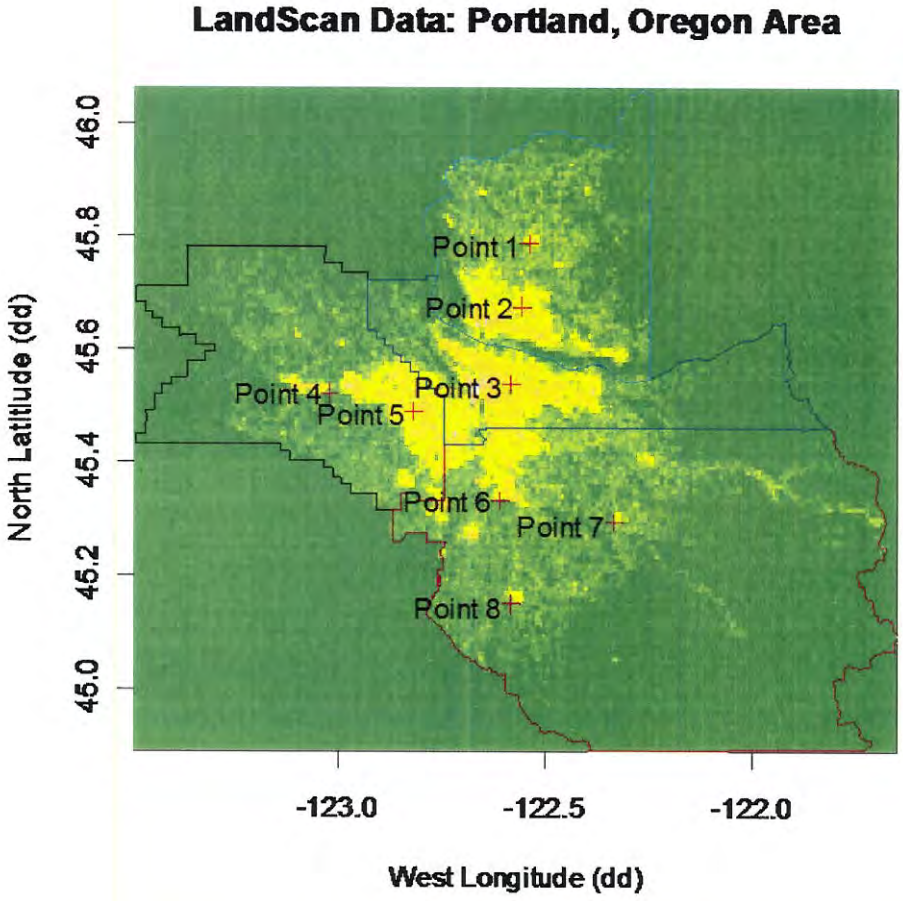


Figure 3-4. LandScan data for the Portland, Oregon area, including the points of initial infections

The data fitting mode proceeds as follows:

We performed a literature search to set all of the disease specific parameters: μ_{ei}, σ_{ei}, μ_{ir}, σ_{ir}. In our case, we used the same values as were used in EpiSims [4].

We also know N (the susceptible population) because we know the population being modeled.

We then proceed to determine fit the key parameters of our model, p_s and p_j, by the following process:

1. An exhaustive search of p_s and p_j to fit the model, where the range of p_s and p_j is [0,1], with a step size of 0.02
2. For each curve generated using the small-world network model, we record
    a. p_s and p_j value
    b. Peak value of I curve
    c. Peak time (days)

3. For each epidemic curve from the full simulation, the peak value and peak time in days is matched with values recorded in the table, according to specific filter
    a. Peak value must be within certain percentage limits (we used 10%)
    b. Peak time must be with certain days (we used 3 days)

    This results in a subset of p_s and p_j values that meet the criteria.
4. From these values of p_s and p_j, we perform a maximum likelihood estimation to choose the best p_s and p_j from the group
5. We then used the p_s and p_j obtained and performed a forecast of the outbreak
6. We visually compared the results against "EpiSims" outbreak (reproduced by Sandia National Laboratories)

3.5.1. Data fitting for Smallpox

Figure 3-5 and Figure 3-6 below for smallpox, I_0 equal to 100 and 1000, show the resulting subset of p_s and p_j values that meet the criteria for smallpox (step 4 in the process above). Note that we are only displaying three out of the eight points in the graph because otherwise the plot is not readable (but the analysis remains valid when all eight points are displayed).

Example Fit: Smallpox, I_0=100

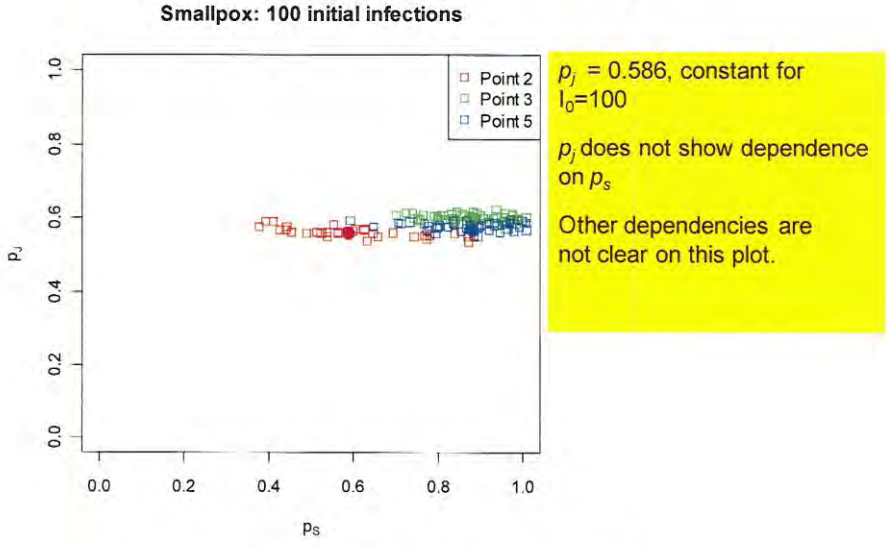


Figure 3-5. Smallpox, I_0=100

Example Fit: Smallpox, I_0=1000

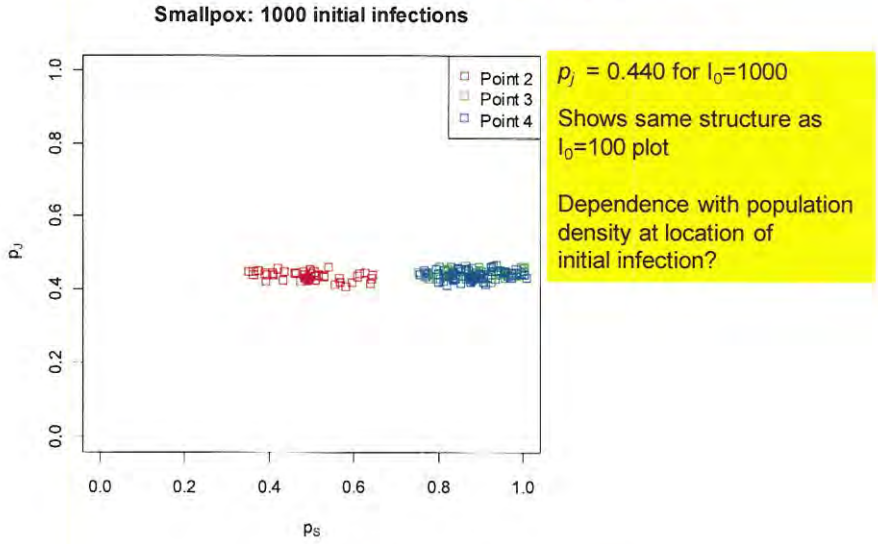


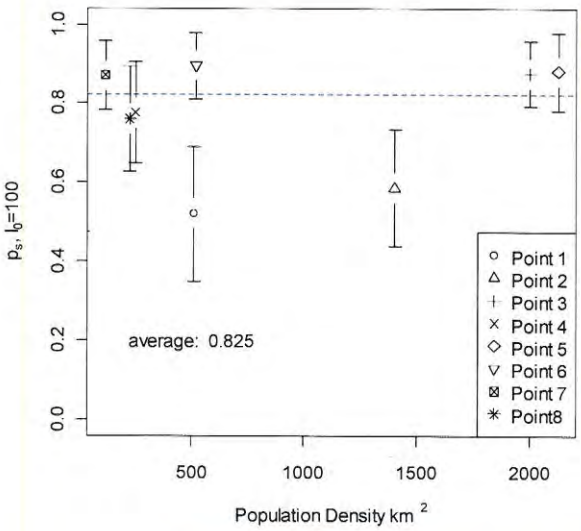
Figure 3-6. Smallpox I_0=1000

From these values, we try to evaluate whether or not the parameters p_s and p_j seem to be affected by location population density. Clearly from these graphs above, p_j is showing no dependence and appears to be constant at 0.440.

Our next question was whether or not there was a formula for a dependence on the parameter p_s on local population density. The plots Figure 3-7 and Figure 3-8 below show the analysis of p_s for all eight points. There did not appear to be a linear relationship between the p_s values and local population density. When the error bars are taken into consideration, it appears that p_s is well described by a constant value of 0.825 for 100 initial infections and 0.721 for 1000 initial infections.

Does P_s Depend on Population Density at Point of Infection For Smallpox?

Smallpox: I_0=100



Local population density at points 1-8 derived from LandScan data

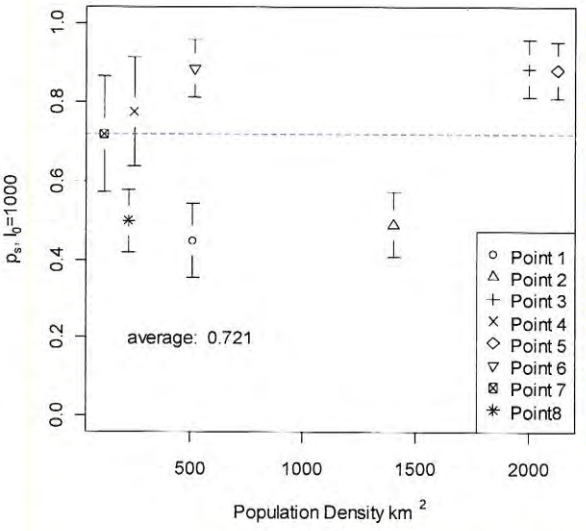
No dependence of p_s on population density

p_s is well described by a constant value = 0.825 for smallpox with 100 initial infected

Figure 3-7. Smallpox, I_0 = 100

Does P_s Depend on Population Density at Point of Infection For Smallpox (Continued)?

Smallpox: I_0=1000



Local population density at points 1-8 derived from LandScan data

No dependence of p_s on population density

p_s is well described by a constant value = 0.721 for smallpox with 1000 initial infected

Figure 3-8. Smallpox, I_0 = 1000

Using these values for p_s and p_j, we wanted to visually validate our model. So we ran our model with those parameters and compared to the original EpiSims data [4] as reproduced by Sandia National Laboratories. Figure 3-9 below shows the results.

Our model fits this complex curve surprisingly well, despite a bit of overshooting magnitude of the epidemic. It captures the complex shape of the multiple waves in the data, something that is not possible in traditional SEIR models.

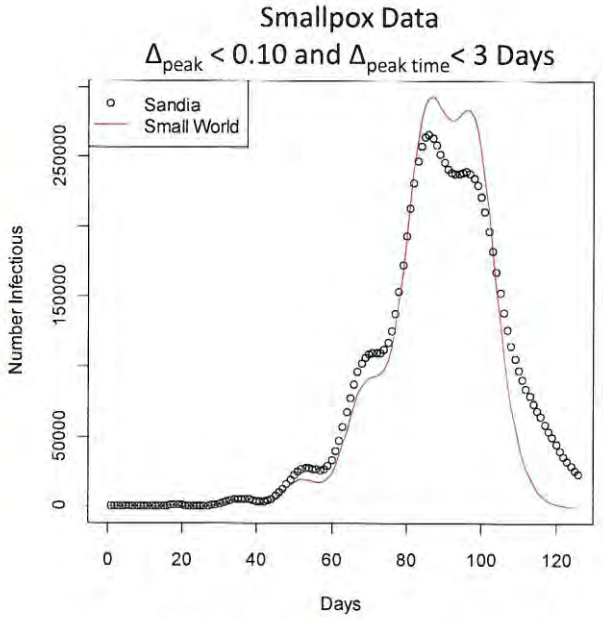
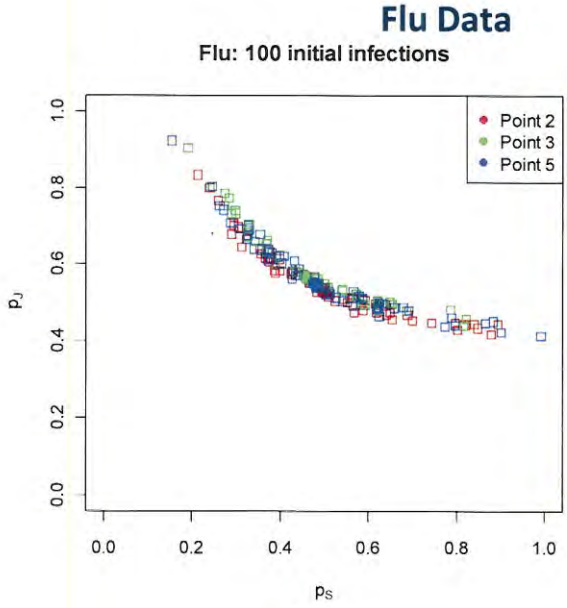


Figure 3-9. Smallpox Data

3.5.2. Data fitting for Influenza (Flu)

Figure 3-10 and Figure 3-11 below show the resulting subset of p_s and p_j values that meet the criteria for influenza (step 4 in the data fitting process on page 10). Note that we are only displaying three out of the eight points in the graph because otherwise the plot is not readable (but the analysis remains valid when all eight points are displayed).

In the case of influenza, it appears that the parameters, p_s and p_j, are interchangeable. This is possibly consistent with the fact that influenza is highly transmissible; thereby it spreads to random contacts just as easily as to near neighbor contacts.



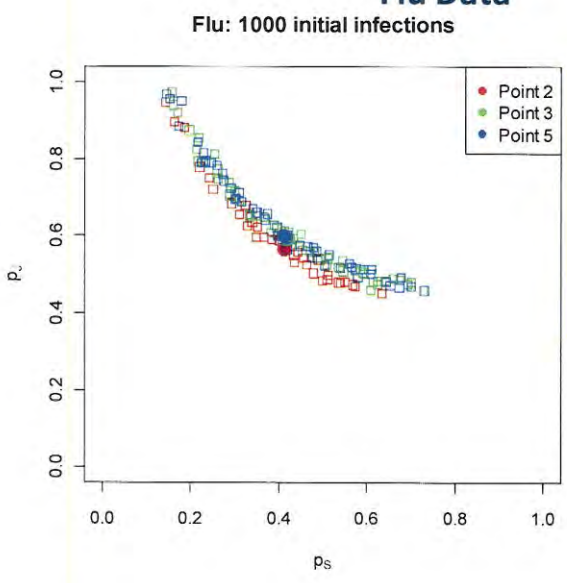
Flu data shows different structure in p_s vs. p_j plots

A range of values in p_s, p_j space provide good fits to the epidemic curves

The median values (shown by the solid circles are roughly constant. For 100 initial infections:

p_s=0.484
p_j= 0.556

Figure 3-10. Flu, I_0 = 100



Insignificant dependence on number initially infected.

Median for 1000 initial infections:

p_s=0.407
p_j= 0.591

Figure 3-11. Flu, I_0=1000

Because no significant dependence on number of initially infected is apparent, we use the median value for the number of initial infections.

For 100 initial infections, $p_s = 0.484$ and $p_j = 0.556$.

For 1000 initial infections, $p_s = 0.407$ and $p_j = 0.591$.

Since we only have two points for I_0, 100 and 1000, we assume a linear relationship between these values to determine p_s and p_j for a given I_0.

Using these values for p_s and p_j, we visually validate our model. Figure 3-12 below shows how our model is an extremely close fit to the EpiSims output (as implemented by Sandia National Laboratories).

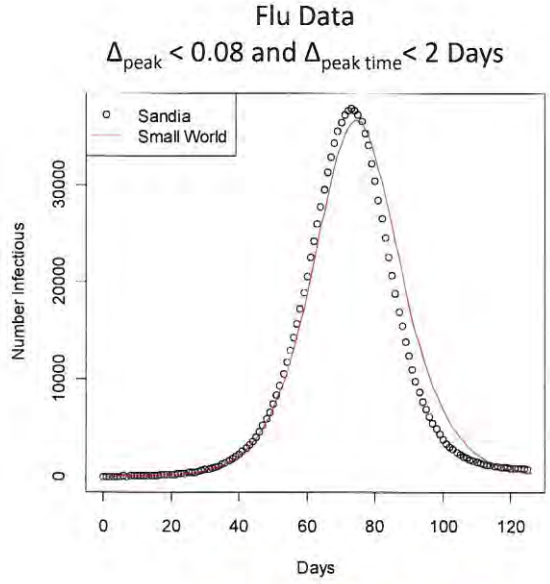


Figure 3-12. Flu Data

3.5.3. Data fitting for Plague

Figure 3-13 and Figure 3-14. Pneumonic Plague I0 = 1000 below, show the resulting subset of p_s and p_j values that meet the criteria for plague (step 4 in the data fitting process on page 10). Note that we are only displaying three out of the eight points in the graph because otherwise the plot is not readable (but the analysis remains valid when all eight points are displayed).

Plague shows yet a different structure where p_j shows a strong dependence on p_s.

We use the media values of p_s and p_j, given by:

$p_s = 0.263$ and $p_j = 0.462$ for I_0=100

$p_s = 0.232$ and $p_j = 0.482$ for I_0=100

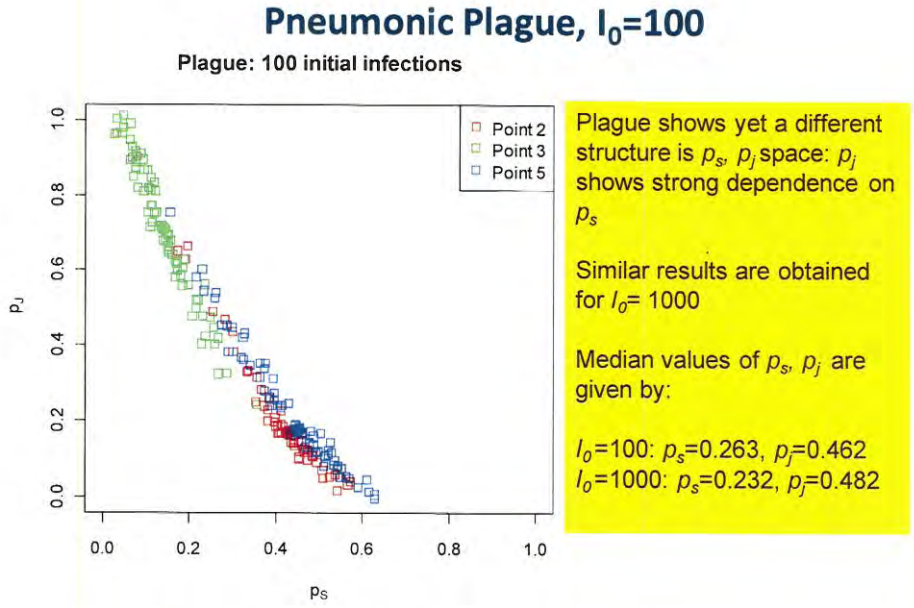


Figure 3-13. Pneumonic Plaque, I_0 = 100

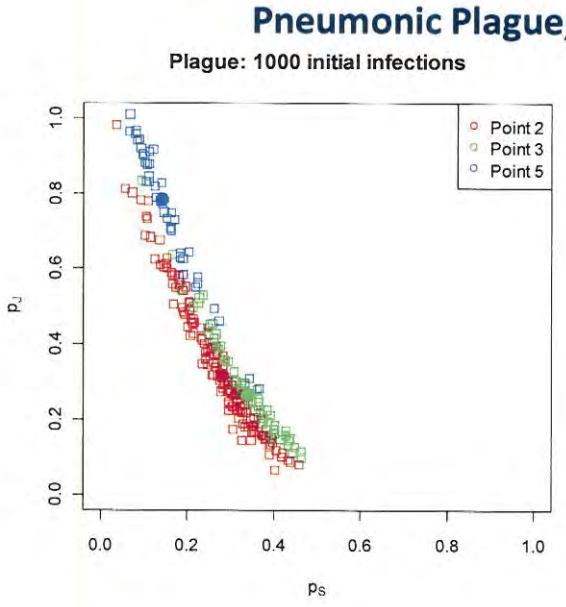


Figure 3-14. Pneumonic Plague I_0 = 1000

Using these values for p_s and p_j, we visually validate our model. Figure 3-15, Below, shows how our model is an extremely close fit to the EpiSims output (as implemented by Sandia National Laboratories).

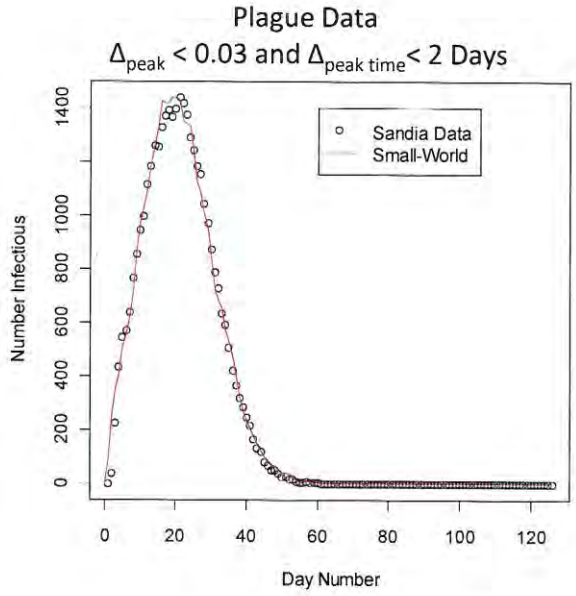


Figure 3-15. Plague Data

3.6. Assumptions and Limitations

Our small world network model is very simple by design. It fits itself to the data given, and scales realistically to different, but related scenarios. Since we are using a simulation, we are not constrained by approximations used in finding analytic solutions to this problem, as in previous studies. However, the models underlying the simulation have the following limitations.

Our small world network model implements a square graph structure, where each node in the square represents a person. Because we have chosen a square, there is an implicit assumption that each person is connected closely, on average, to four people. Close connections generally represent family members, and possibly co-workers. As discussed above, the average family size is actually 3.14 , so in the future, we could break links in our square to more closely represent this actual figure for family size to determine whether or not the fitting is improved.

Another assumption we make when fitting our model to data, is that the disease behaves the same at different time periods of the year. The problem with addressing this limitation is mostly one of limited data available. If the disease behaved differently in different time periods of the year, it could be caused by various factors such as environmental changes that enable more or less aggressive disease transmissions, changes in peoples' behaviors (for example, if people stay home more in the winter because it is too cold), or changes in the disease's ability to survive in different temperatures. These factors would have to be known and then included in the model.

The model takes as input a susceptible population number, *N*. The model does not have any spatial awareness built into it. It returns the number of people that will be infected on each day. The model does not currently specify which people a simulation should select to infect (however, the simulation that uses the model is free to implement whatever heuristic or model it chooses to).

3.7. Model Uncertainties

The small world network model is a simulation with a stochastic element. Therefore, there is uncertainty in each run of the model. In fitting our models to data, we generally averaged ten instantiations of the small world network model. Our studies of the model have shown that there is variation between runs; however, they are not very significant.

3.8. Recommendations for Future Work

- Experiment with changing the number of nearest neighbors by randomly breaking links in our underlying square network structure
- Fit model to real outbreaks
- Do further studies on relationship of model parameters to number of initially infected
- Do further studies on relationship of model parameters p_s and p_j to density of initial infected

4. Summary

The secondary infection small world network model is a fast running simulation that can be incorporated into external programs to return

- The number of people in the Susceptible [S] compartment on a daily basis
- The number of people in the Exposed [E] compartment on a daily basis
- The number of people in the Infectious [I] compartment on a daily basis
- The number of people in the Recovered [R] compartment on a daily basis
- Maximum day of the outbreak
- The duration of the outbreak
- The number of total infections throughout the outbreak

In this report we have described the benefits of the small world network model and how we fit our models to a complex simulation model created by Sandia National Laboratories based on EpiSims [4] in the Portland, Oregon area.

5. Referenced Documents

[1] Bailey, N. (1975) "The Mathematical Theory of Infectious Diseases and its Applications" Charles Griffin & Company, London.

[2] Anderson, M. & May, M. (1979) "Population Biology of Infectious Diseases: Part I" Nature, 280, 361.

[3] Hethcote, H. (2000) "The Mathematics of Infectious Diseases," SIAM Review, 42(4), 599.

[4] Eubank, S. et al. (2004) "Modeling disease outbreaks in realistic urban social networks", Nature, 429, 180.

[5] Saramaki, J, & Kaski, K. (2005) "Modeling development of epidemics with dynamic small-world networks" J. Theor. Bio., 234, 413-421.

DEPARTMENT OF DEFENSE

DEFENSE TECHNICAL
INFORMATION CENTER
8725 JOHN J. KINGMAN ROAD,
SUITE 0944
FT. BELVOIR, VA 22060-6201
2 CYS ATTN: DTIC/OCA

DEPARTMENT OF DEFENSE CONTRACTORS

ITT INDUSTRIES
ITT SYSTEMS CORPORATION
1680 TEXAS STREET, SE
KIRTLAND AFB, NM 87117-5669
2 CYS ATTN: DTRIAC

APPLIED RESEARCH ASSOCIATES,
INC.
801 N. QUINCY STREET
ARLINGTON, VA 22203
ATTN: G. REEVES